MEME16203Linear Models

Assignment 1

UNIVERSITI TUNKU ABDUL RAHMAN

Faculty:	FES	Unit Code:	MEME16603
Course:	MAC	Unit Title:	Generalized Linear Models
Year:	1,2	Lecturer:	Dr Yong Chin Khian
Session:	September 2024		
Due by:	9/11/2024		

Q1. You are given the following probability density function for a single random variable, X:

$$f(x) = \left(\frac{\theta}{2\pi x^3}\right)^{\frac{1}{2}} \exp\left(-\frac{\theta(x-1)^2}{2x}\right).$$

(a) Find the score function, U.

(b) Find the information matrix, $\mathcal{J}$.

(c) Find the asymtotic distribution of $\frac{U}{\sqrt{\mathcal{J}}}$.

(3 marks)

Q2. You are given the following probability density function for a single random variable, X:

$$f(x) = \frac{\alpha\theta^\alpha}{(x+\theta)^{\alpha+1}}.$$

(a) Find $E\left(\frac{1}{x+\theta}\right)$ and $E\left(\frac{1}{(x+\theta)^2}\right)$.

(b) Find the score function, $\mathbf{U} = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}$.

(c) Find the information matrix, $\mathcal{J}$.

(3 marks)

Q3. You are given the following probability density function for a single random variable, X:

$$f(x) = \left(\frac{2}{2\pi x^3}\right)^{\frac{1}{2}} \exp\left(-\frac{2z^2}{2x}\right), z = \frac{x-\mu}{\mu}.$$

(a) Find the score function, U.

(b) Find the information matrix, $\mathcal{J}$.

(2 marks)

Q4. The following logistic model for the probability of passing an exam is fitted:

Response variable: Probability of passing the exam Response distribution: Binomial Parameter	β
Intercept	-1.21
Familiar with R	
Yes	1
No	0
Hours of studying the exam material	0.0053
Score on SPM Additional Mathematics	
4 or less	-3.38
5	0
6	1.24
7 or more	1.58

Calculate the probability and odds of passing the exam for someone who is not famaliar with R, has studied the exam material for 171 hours, and who passed SPM Additional Mathematics with a 6,

(a) if the link selected is logit,

(b) if the link selected is probit,

(c) if the selected link is complementary log-log.

(3 marks)

Q5. You are given the following results for two generalized linear models fit to the same data:

Model	AIC
$g(\pi) = \beta_0$	89.2
$g(\pi) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$	88.4

Calculate the likelihood ratio statistic to test $\beta_1 = \beta_2 = \beta_3 = 0$. (1 mark)

Q6. The compressive strength of an alloy fastener use in aircraftconstruction is being studied. Ten loads were selected over theb range 2500-4300 psi and a number of fasteners were tested at those loads. The numbers of fasteners failing at each load were recorded. The data for the first 3 observations are shown below.

Load, x(psi)	2500	2700	2900	$\cdots$
Sample size, n	50	70	100	$\cdots$
Number Failing, y	10	17	30	$\cdots$

The R output is given below.

```
glm(formula = failing ~ x, family = binomial(link = "logit"))
```

```
Deviance Residuals:
   Min      1Q  Median      3Q     Max
-4.139  -3.457  -0.032   2.398   4.918

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.0801749  0.1082146  -0.741    0.459
x            0.0006361  0.0014470   0.440    0.660

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 112.83  on 9  degrees of freedom
Residual deviance: 112.64  on 8  degrees of freedom
AIC: 161.36

Number of Fisher Scoring iterations: 3
```

(a) Fit a logistic regression model to the data. Use simple linear regression model as the structure for the linear predictor.

(b) Does the model deviance indicare that the the logistic regression from part (a) is adequate.

(c) Find the estimated probability of number of fasteners failing for a load with 3500 psi and a sample size of 85.

(d) Interpret the slope β_2.

(e) Find the diviance residual for the first observation, d_1.

(5 marks)

Q7. A binary response is modeled with a generalized linear model and a probit link. The fitted model is

$$g(\pi) = 0.27 + 0.49x_1 + 0.64x_2$$

Determine the probabilty of an event when $x_1 = 2$ ans $x_2 = 3$. (1 mark)

Q8. The regression model $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \epsilon$ is being investigated. The following maximized log-likelihoods are obtained:

- Using only intercept term:-1126.91
- Using only intercept term, X_1 and X_2: -1122.41
- Using all four terms: -1121.91

The null hypothesis $\beta_1 = \beta_2 = \beta_2 = 0$ is being tested using the likelihood ratio test. Determine the samallest significance level at which you reject the null hypothesis. (1 mark)

MEME16203Linear Models

Q9. For a generalized linear model, the response didtribution is binomial. For a cell with observed value 1, the fitted probability is 0.223. Calculate the Pearson residual for that cell. (1 mark)